Ultrafast 2D-IR spectroscopy of haemoproteins

N. SIMPSON, N. T. HUNT*

Department of Physics, University of Strathclyde, SUPA, 107 Rottenrow East, Glasgow, G4 0NG, UK

* Corresponding author email address: neil.hunt@strath.ac.uk

Applications of ultrafast two-dimensional infrared (2D-IR) spectroscopy to study the structural dynamics of haem-containing proteins are reviewed. The 2D-IR experiments discussed exploit diatomic ligands bound to the haem as reporters on the dynamic protein environment in the electronic ground-state. This is possible because fluctuations of the protein give rise to inhomogeneous broadening of the ligand stretching vibrational mode that is manifest as spectral diffusion in a time-resolved 2D-IR measurement. Methods for measuring and quantifying spectral diffusion data are introduced, prior to a discussion of recent results focussing on the influence of protein structure, water ingress into the haem pocket and substrate binding on the measured dynamics. Particular emphasis will be placed on proteins featuring the ferric oxidation state of the haem ligated by a nitric oxide molecule, though comparisons with other haem systems will be drawn throughout.

Introduction

The haemoproteins are ubiquitous in biology, where they perform a range of physiologically-important roles [1], including ligand and electron transport [2, 3], signalling [4, 5], and substrate oxidation [6]. As a family, the haemoproteins have been the subject of much enquiry. In many cases, molecular structures have been obtained via X-ray crystallography [7, 8] and NMR spectroscopy [9-11], amongst other methods, while molecular and biochemical approaches have determined their function [12]. Indeed, our comprehension has progressed to the extent that haem-derived systems are now important templates for synthetic biology and biomimetic molecular systems [13-16]. Despite these advances however, many questions still remain regarding the way in which the two components of structure and function are linked at the molecular level.

One of the most interesting facets of the haemoproteins, and one that is directly relevant to their use in *de novo* synthetic applications, is that, despite their biological versatility, they display a number of recurring structural attributes and conserved features near the haem. For example, proximal and distal histidine ligands are highly conserved amongst globin and peroxidase sub-families [17-19] while as few as two point mutations have been shown to be sufficient to drastically alter the functionality of a given haem protein [13]. Overall, this suggests that the influence of structural features in the vicinity of the haem upon function is subtle and complex. Further modulation of haemoprotein behaviour arises from physical aspects of the wider protein environment surrounding the haem site, such as residues or secondary structure elements that control solvent accessibility [3] or substrate binding [20, 21]. Ultimately, this means that, at the molecular-level, haem protein function is an amalgamation of highly localised, residue-specific interactions in the active site, controlling the atomistic details of a chemical process, with the broader properties of more distant regions of the amino acid sequence that control substrate access.

One aspect of this multi-component problem that we do not yet fully comprehend is the impact of dynamic motion and, in particular, ultrafast dynamics on enzyme function. Protein structures display motions on a range of timescales spanning femtoseconds to seconds, many of which have the potential to be functionally relevant [22]. At the faster end of this range lie the ultrafast (fs-ps) processes that are linked to important events such as those that give access to a transition state or which underpin or initiate slower changes in molecular structure. Our understanding of these processes is currently lacking because effective measurement techniques have paralleled the development of the ultrafast laser and so emerged only relatively recently in comparison to e.g. X-ray crystallography. The result is that our knowledge of how dynamic phenomena such as protein vibrational motion [23-28], conformational switching [29, 30] or hydrogen bond exchange [31, 32] influence biomolecule behaviour is incomplete. It is especially noteworthy that the ubiquitous biological solvent, water, features a diverse range of dynamics in the ultrafast domain and our ability to comprehend the biological involvement of so-called non-structural water is therefore similarly limited [33-37].

Considerable strides towards providing information on haemoprotein dynamics have been made by ultrafast spectroscopic techniques that exploit the rich electronic absorption spectrum of the haem moiety. Pump-probe or Raman-based methods using UV or visible wavelength excitation of haem electronic transitions have investigated processes such as ligand binding to the haem and the impact of haem group vibrations [38-44]. Linking such information with an understanding of the wider protein dynamics under equilibrium, or electronic ground state, conditions is however desirable and the technique of 2D-IR spectroscopy, on which we will focus here, has become established as a particularly powerful tool in this respect [45-52].

The first application of 2D-IR spectroscopy to haem proteins was carried out by the Fayer group using multidimensional vibrational echo spectroscopy to establish the dephasing dynamics of a carbon monoxide (CO) ligand bound to the haem centre of myoglobin (Mb) [53]. This built upon several studies using the 1D vibrational echo method on which the 2D-IR technique is based and together these laid the groundwork for the studies featured in this review [54-60]. The fact that the frequency of the stretching vibrational mode of the CO ligand had been shown previously to be sensitive to fluctuations in its local environment formed the basis of the experiment. This sensitivity was manifest as line broadening in linear IR absorption experiments [61, 62], but the use of a three pulse sequence in the vibrational echo measurement allowed information on the underlying dephasing and spectral diffusion dynamics causing the broadening, and so the dynamics of the protein environment itself, to be elucidated. There followed a steady evolution of the experimental and theoretical methodology toward the current state-of-the art, which visualises the dynamics of haem proteins through the 2D-IR lineshape and which has been widely adopted [48, 50].

Since these early experiments, the range of haem systems studied has undergone considerable expansion, not only in terms of the identity of the proteins themselves but also the nature of the diatomic reporter ligand, the oxidation state of the ferric centre and incorporation of substrates. This means that we are now in a position to draw these results together and so address the broader picture of how dynamics influence function and the aim of this article is to embark upon this process. The text is structured as follows: The next section will introduce the 2D-IR experiments used to extract haem protein dynamics while the following section will review recent progress. In particular we will focus upon work on nitrosylated ferric haem proteins, but will also draw comparisons across the literature as appropriate.

2D-IR spectral diffusion measurements

The technique of 2D-IR spectroscopy has been reviewed many times in recent years and is now the topic of a text-book [45-52]. As a result, a full discussion of the methodology is superfluous here but a brief overview will be provided prior to a more detailed consideration of the specific 2D-IR experiment used to measure spectral diffusion, and so haemoprotein structural dynamics.

The 2D-IR experiment uses ultrafast infrared lasers to produce a 2D-spectrum in the IR region analogous to that of the 2D-NMR methods based on the spin echo technique that have revolutionised structural biology [63]. Briefly, a 2D-IR spectrum is obtained from a 3^{rd} order nonlinear optical process in which three resonant interactions occur between ultra-short IR laser pulses and vibrational transitions of the sample, leading to the emission of a 4^{th} signal pulse or echo. A diagram of the experimental pulse sequence is shown in Fig 1(a) where the time delays between successive pulses are represented by τ, T_w and t respectively. Various experimental methods exist to obtain 2D-IR spectra but all produce a correlation map of the excitation (pump) frequency, *i.e.* the frequency of the first laser-sample interaction, and the detection (probe) frequency, *i.e.* the frequency of the third interaction [45-52]. Analogously to 2D-NMR measurements, this results in the peaks visible in a linear IR absorption spectrum of the system appearing on the diagonal of the 2D plot, with off-diagonal peaks indicating the existence of vibrational couplings between modes [45-52].

The main advantage of 2D-IR over NMR however lies in an intrinsic temporal resolution that is on the order of ~100 fs. Typically, this is exploited by observing the evolution of the 2D-IR spectrum of a system as a function of the waiting time (between pulses 2 and 3, labelled T_w in Fig 1(a)). Experiments recorded in this manner provide access to ultrafast dynamics arising from processes such as vibrational energy transfer or chemical exchange, which give rise to time-resolved changes in off-diagonal peak amplitudes in the spectrum [45-52]. In the work

that follows, we will focus not upon off-diagonal peaks but upon the evolution in 2D shape of diagonal peaks in a 2D-IR spectrum as T_w changes, beginning with a review of how this so-called spectral diffusion effect can be used to directly measure the structural dynamics of haem proteins.

In all of the studies that will be discussed below, the basic tenet is the location of a vibrational reporter group in the haem pocket. In the majority of cases, this reporter group takes the form of a diatomic species bound directly to the Fe centre of the haem. As mentioned briefly above, CO, which coordinates strongly to ferrous (Fe(II)) haem centres, is to date, the most widely-used reporter species for these applications [61]. This is largely due to the fact that the stretching vibrational mode of haem-bound CO (ν_{CO}) is located at a frequency of ~1900-2000 cm^{-1}, meaning that it is not coincident with intense infrared absorptions due to either the solvent or protein backbone. Furthermore, the ν_{CO} absorption is of sufficient intensity as to be resolvable when the protein is present at the few-millimolar (mM) to sub-mM concentrations in aqueous (or more often deuterated water) solution that are typical; this limit often being dictated by protein stability considerations. Another ligand that has been exploited for haem protein studies of this nature is nitric oxide (NO). This ligand displays a somewhat weaker stretching vibrational mode than CO, but has the advantage of being able to bind to haem proteins in both the ferrous and ferric (Fe(III)) oxidation states. As a result, NO has typically been used to study ferric haems, which display a ν_{NO} mode located near 1850-1900 cm^{-1}. In contrast, the corresponding mode of the ferrous form is located at a lower frequency (~1600 cm^{-1}) and is often obscured by protein backbone absorptions [61]. More recently, applications using other reporter groups such as thio- or seleno-cyanate have shown considerable promise for ultrafast spectroscopy of haem proteins [64].

The stretching vibrational mode of all of these reporter groups has been shown to display inhomogeneous broadening in the protein or solvent environment. This arises because the

mode frequency is sensitive to changes in the covalent and non-covalent interactions of the reporter group with the Fe centre and local protein residue side chains or solvent molecules respectively. In addition, changes in the general electrostatic environment due to the protein scaffold influence the stretching frequency of the reporter group via a more indirect mechanism that has been described as a Stark effect-type interaction [62, 65]. In a given sample therefore, each reporter group occupies one of a continuum of substates caused by fluctuations of the protein structure. More formally, the vibrational transition of the reporter is described as being coupled to the low frequency density of states of the solvent and protein. This means that when a linear IR absorption experiment is performed, an ensemble average over all possible environments accessible to the reporter ligand is measured and a broadened Gaussian lineshape results. This is shown schematically in Fig 1(b).

Such a measurement does not however reflect the dynamic environment of the reporter. In solution, or in the haem pocket of a protein, the respective solvent motion or protein structural dynamics, leads to time-dependent variations in vibrational frequency of the reporter. As discussed above, in contrast to linear IR experiments, a 2D-IR experiment spreads the infrared absorption of the diatomic haem ligand over a second frequency axis in a manner that can be described as a correlation of excitation frequency with detection frequency [45-52, 66]. Alternatively, the 2D-IR experiment can be thought of as a pump-probe experiment that is frequency resolved in both pump and probe directions. Of particular relevance here, such a methodology introduces a temporal component into the measurement through control of the waiting time (T_w). This is the time between pump (excitation) and probe (detection) events and allows capture of the vibrational dynamics that are not discerned in a linear absorption experiment.

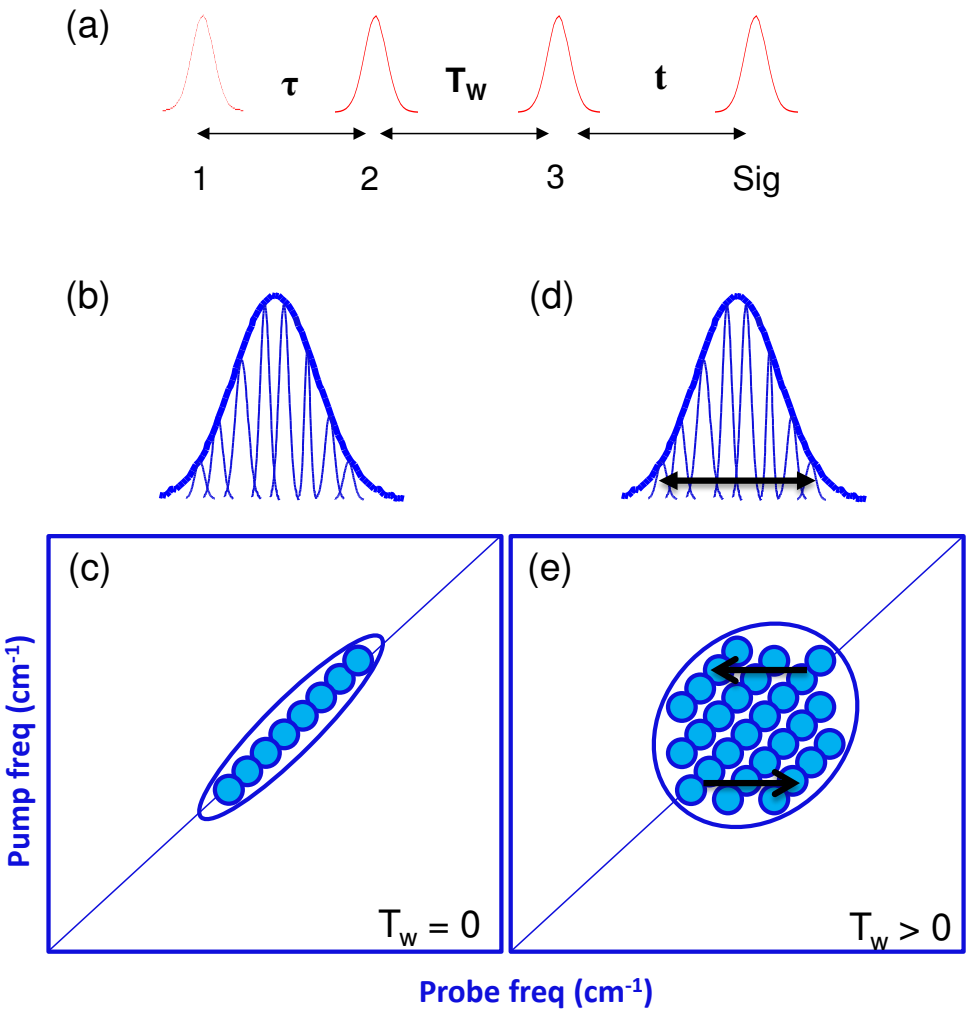

Figure 1: *a) Diagram of a typical laser pulse sequence used to generate a 2D-IR spectrum. b-e) Schematic diagram of the effect of inhomogeneous broadening on linear and 2D-IR spectra of a vibrational reporter group bound to a haem protein. b) and d) display model linear absorption spectra while c) and e) show accompanying 2D-IR spectra at short (c) and long (e) waiting times. Black arrows indicate the effect of spectral diffusion dynamics arising from fluctuations of the reporter group environment on each type of spectrum.*

A 2D-IR experiment on a reporter group bound to a haem centre thus allows investigation of the dynamics underlying the inhomogeneous broadening observed in a linear IR absorption

experiment: The pump pulse in a 2D-IR experiment effectively labels a sub-population of the ensemble of reporter ligands by exciting only those with specific instantaneous frequencies within the continuum. If the probe pulse arrives in the sample at a very short time after the pump ($T_w \sim 0$), i.e. at a time delay shorter than the underlying dynamics of the system (e.g. water motion, protein fluctuation), then the frequency of the reporter mode does not change between excitation and detection. This means that the pump and probe frequency are the same and a diagonal response is found in the 2D-IR spectrum (see Fig 1(c)). The result is that, if the waiting time in the 2D-IR experiment is short on the timescale of the processes causing inhomogeneous broadening, then the 2D-IR lineshape will be diagonally elongated (Fig 1(c)); when probed, the sample retains a 'memory' of the state it was in when excited and a correlation exists between the pump and probe frequency.

If the waiting time in the 2D-IR experiment is increased to the point where it is comparable to the underlying dynamics of the system ($T_w>0$), then the environment of an excited reporter molecule will change during the time that elapses between excitation and detection. In this case, the pump frequency and probe frequency will differ (Fig 1(d, e)). This causes an anti-diagonal broadening of the 2D lineshape leading to a progressively more circular profile as the waiting time increases. This lineshape evolution directly mirrors the changing frequency of the reporter vibrational mode as it samples the entire continuum of environments available to it between the pump and probe events. Thus, dynamic motion causes the memory of the initial state of the system when excited to be lost and the correlation between pump and probe frequency is not retained (Fig 1(e)). This effect is termed spectral diffusion and will occur providing the dynamics influence either the covalent or non-covalent bonding to the reporter group or change the electrostatic environment that is experienced by it [45-52]. As such, these reporter groups are widely considered to sample their local environment, though we will also see below that some longer range effects can be sampled by this measurement.

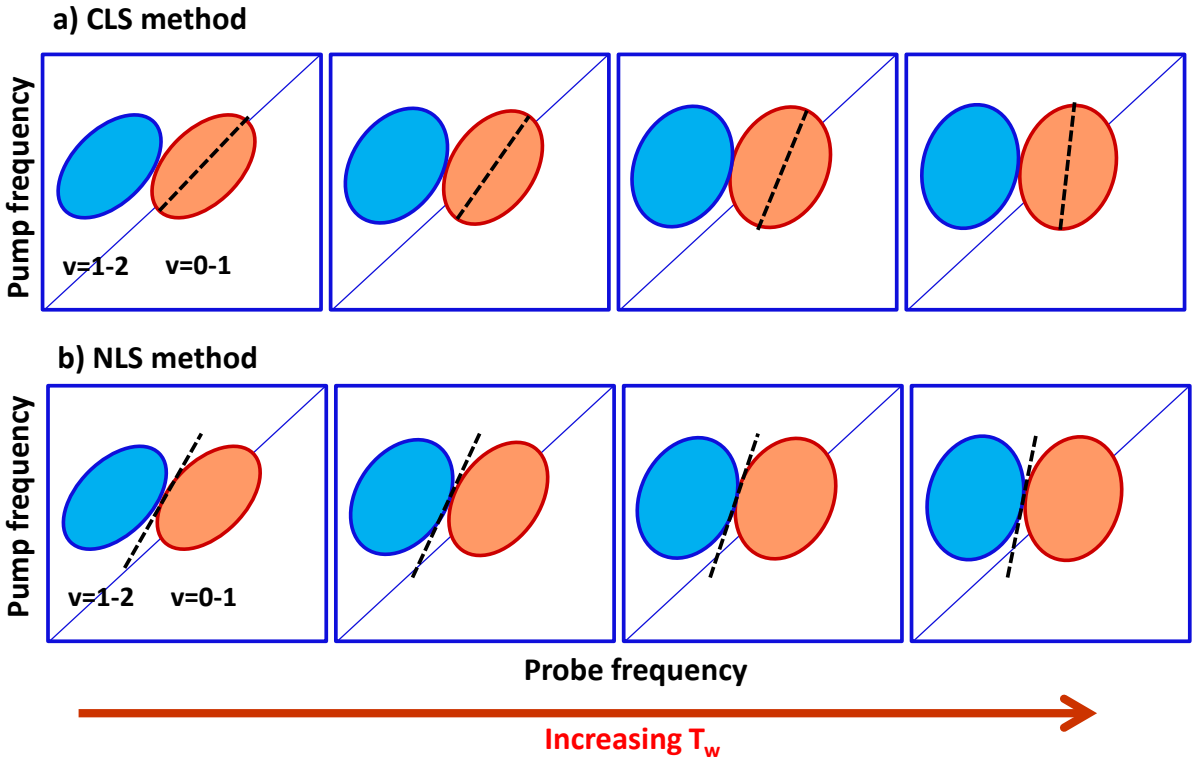

Figure 2: *Schematic diagram of a) Centre line Slope (CLS) and b) Nodal Line Slope (NLS) methods for quantifying spectral diffusion dynamics from a 2D-IR measurement. Orange and blue represent components of the spectrum of opposite phase corresponding to the v=0-1 and v=1-2 transitions of the* ν_{XO} *(X=N,C) mode of the reporter ligand.*

The objective for a 2D-IR experiment is thus to determine the frequency-frequency correlation function FFCF of the reporter vibrational mode,

$$C(t) = \langle \delta\omega(t)\delta\omega(0) \rangle \qquad (1)$$

which provides a measurement of the frequency and amplitude of the equilibrium fluctuations underpinning the change in lineshape. Extracting quantitative data relating to the FFCF from the acquired 2D-IR spectra has been achieved in a variety of ways that have been shown to be largely equivalent and to accurately reflect the FFCF dynamics [26, 34, 67, 68]. These

include the centre line slope method (CLS) in which the line corresponding to the highest point of the 2D-IR signal at each pump frequency is identified (Fig 2(a)) and the inverse of the gradient of this line is plotted vs T_w (Fig 3) [67]. Alternatively, the nodal line slope method (NLS) utilises the fact that the 2D-IR peak from the reporter ligand consists of two parts due to the v=0-1 and v=1-2 transitions of the stretching vibrational mode. These appear in the 2D-IR spectrum with opposite amplitude (phase). The NLS method identifies the point at which the signal changes sign for each pump frequency, the nodal line (Fig 2(b)), and plots the inverse gradient of this line as a function of T_w (Fig 3) [69]. Two other methods include plotting the elipticity parameter of the lineshape versus T_w [68] or fitting to a 2D-Gaussian lineshape to extract a correlation parameter which can be plotted as a function of T_w [26].

Irrespective of the method employed, the spectral diffusion dynamics extracted have a similar form (Fig 3). These are invariably fit using an exponential decay function, though the number of timescales included depends upon the individual system. The main region of interest focuses on the values of T_w where the lineshape is changing in this exponential manner (exponential region, Fig 3) reflecting the dynamic loss of pump-probe frequency correlation. The timescales of the individual contributions reflect the dynamic processes influencing spectral diffusion and provide a measurement of the dynamics of the FFCF of the reporter mode.

Additional information can be gleaned from the value of the parameter at short pump-probe delay times (value at T_w=0, Fig 3). In a fully inhomogeneous system, with complete correlation of pump and probe frequency at the shortest measurement time, this value would be unity. Realistically however, this value is reduced from one at short times by contributions from the homogeneous linewidth of the transition being studied, which contributes to the

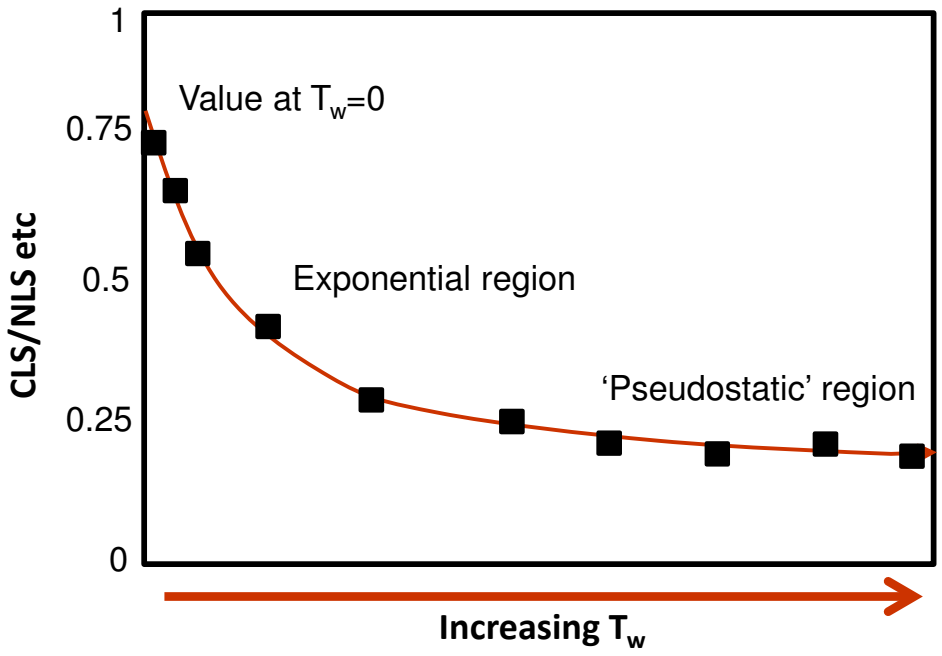

Figure 3: *Schematic representation of the results obtained from a CLS or NLS analysis of a 2D-IR spectral diffusion experiment, showing typical features observed in studies of haemoprotein samples.*

anti-diagonal linewidth of the 2D-IR signal at short T_w, and dynamics that influence the inhomogeneous broadening but which occur faster than the time resolution of the experiment (typically 100 fs).

Similarly, the value of the spectral diffusion parameter does not always fall to zero within the times accessible by the experiment (pseudostatic region, Fig 3). This residual value of the spectral diffusion parameter indicates that some of the processes contributing to inhomogeneous broadening are not complete by the maximum timescale of the experiment, which is defined by the vibrational relaxation time of the stretching vibration of the reporter ligand. For a NO ligand, this is typically of the order of 15-25 ps. A significant pseudostatic contribution is often observed in haemoprotein systems where slower components of the structural dynamics occur on a range of timescales. Thus, by combining 2D-IR spectroscopy of a diatomic reporter ligand bound to the haem centre of a protein with lineshape analysis, a

detailed appreciation of the structural and solvent dynamics that affect the haem ligand can be obtained.

The role of distal cavity residues

As mentioned above, the earliest applications of spectral diffusion to understand the dynamics of haemoproteins were performed on myoglobin. Biologically, Mb is responsible for storage and transport of molecular oxygen in muscle tissue; a function it performs in the ferrous oxidation state. As such, the use of carbonmomoxy ligand to probe ferrous Mb was a logical choice from a physiological perspective. By contrast, ferric Mb occurs naturally only as a minority species and its biological role is less well-defined. Ferric Mb has however been implicated in the scavenging of NO to prevent nitrosative stress and this motivated the first exploitation of the NO probe for 2D-IR applications [70, 71]. A 2D-IR study of the dynamics of NO bound to ferric Mb thus achieved the twin goals of specifically exploring the ferric, nitrosylated form of Mb as well as providing the opportunity to understand the effect of a change of oxidation state on haemoprotein structural dynamics [26, 72].

The IR absorption spectrum of ferric Mb-NO (from horse heart) was found to show a single absorption band attributable to the ν_{NO} stretching vibration [72]. This suggested that only one structural conformation of the NO ligand was present within the haem pocket [61]. The first measurement of the spectral diffusion dynamics of nitrosylated Mb using 2D-IR methods showed that this band was inhomogeneously broadened and that the underlying spectral diffusion dynamics were dominated by a 5 ps exponential relaxation process [72]. This was accompanied by a significant pseudostatic component present in the data at long T_w values (~50 ps) indicating that a substantial contribution to the inhomogeneous broadening of the NO stretching absorption was attributable to slower dynamic processes.

The 2D-IR methodology employed in this initial study was based on double resonance 2D-IR, which employs a frequency-domain acquisition method based on scanning the frequency of a narrow bandwidth pump pulse [48, 50]. This method limits the time resolution of the technique to 1-2 ps and these results were later revisited employing the time domain Fourier Transform (FT) 2D-IR methodology [26]. By replacing the narrow bandwidth pump pulse with an interferometric broadband approach, the FT2D-IR method preserves the 100 fs time resolution afforded by the duration of the ultrafast IR pulses. Through this approach it was determined that a more accurate value of the exponential decay timescale for Mb-NO was 3 ps, though the large pseudostatic component remained, as would be anticipated [26].

Having established the dynamics sensed by the ligand located in the haem pocket of the protein, a significant challenge remained in assigning them to particular aspects of the haem pocket structure and so to establish an understanding of the results that inform our view of protein behaviour. The distal histidine residue is known from crystallography to interact strongly with the haem ligand in Mb via a nitrogen atom located on the histidine ring (Fig 4a) and so might be expected to influence the spectral diffusion dynamics of the haem ligand. Thus a logical step involved the exchange of this residue for a different amino acid, yielding an altered interaction with the reporter ligand. This motivated a comparison of the spectral diffusion dynamics of wild-type ferric Mb-NO with the single point-mutant H64Q [26].

The H64Q mutant of Mb substitutes the histidine group at the distal position with a glutamine residue (Fig 4a), mimicking the distal pocket structure that occurs naturally in a subset of Mb proteins found in elephants and sharks. In contrast to the wild-type Mb sample, the NO stretching absorption of H64Q Mb was found to have two contributions, indicating the presence of two different haem pocket conformations (Fig 4b) [26]. One of these showed a similar ν_{NO} absorption frequency to that of wild type Mb-NO while the more intense of the two bands was shifted some 20cm^{-1} to lower frequency. The glutamine side chain present in

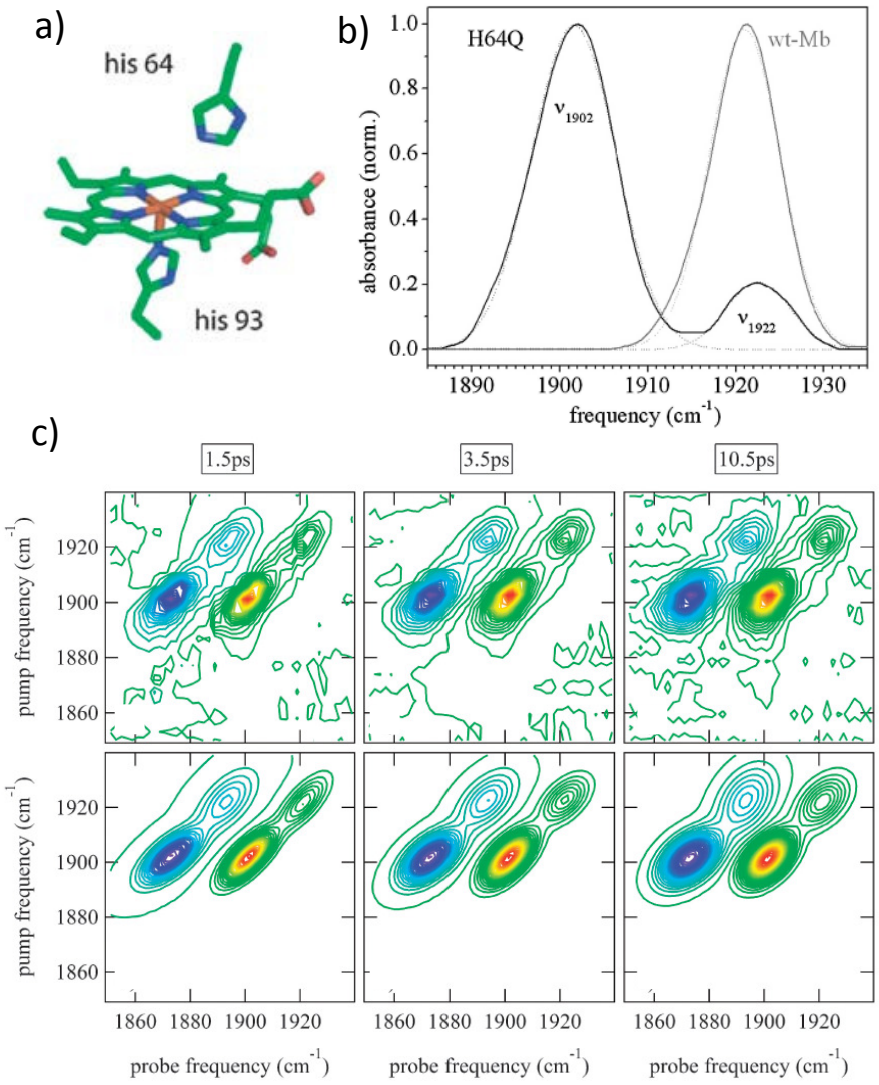

Figure 4: *a) Structure of the haem pocket and distal side chain residue in respect to the ferric haem centre in Mb (PDB code 1YMB) [73]. In the H64Q mutant the His64 residue is replaced with a glutamine. b) FTIR spectra of ferric wt-Mb–NO (grey) and H64Q-NO (black). Solid lines show background-subtracted data, whereas dashed curves show the results of fitting to Gaussian lineshape functions. c) 2D-IR spectra (top row) of NO bound to H64Q recorded at various waiting times,* T_w*. Bottom row shows fits of the experimental data using a two dimensional Gaussian function [26].*

the H64Q mutant is of comparable size to that of histidine and can also offer a similar lone pair interaction with the ligand via a nitrogen atom. By contrast however, glutamine has a greater degree of flexibility than histidine and so the high frequency ν_{NO} band was assigned to

a ‘wild-type-like’ conformation in which the glutamine side chain interacted directly with the NO ligand through a similar lone-pair donation mechanism to that found between the wild-type histidine side chain and the NO. The lower frequency band was assigned to a conformation in which the flexible glutamine side chain was positioned in such a way as to remove this interaction [61].

The results of the 2D-IR spectral diffusion analysis showed that substantial variations were found between the dynamics of the two conformational substates of the H64Q mutation (Fig 4c). The wild-type-like H64Q substate showed very similar dynamics to the wild-type protein, with spectral diffusion values of 3.5 and 3.3ps recovered respectively. By contrast, a longer spectral diffusion time of 9.7ps was observed for the lower frequency substate of the mutant that lacked the direct distal residue interaction [26]. This result suggests that the direct contact between the ligand and the distal haem scaffold, whether via histidine or glutamine, is apparently linked to the observation of spectral diffusion time-scales on the order of 3 ps. In addition, this result gives some insight as to the nature of the protein environment. Each substate giving rise to a separate NO absorption band can be thought of as a local minimum in the potential surface associated with the protein structure. The spectral diffusion dynamics of each mode then reports on structural fluctuations within that specific local minimum. Though not observed in this case, examples exist where 2D-IR spectroscopy can also observe transitions between minima, known as conformational switching *vide infra* [29].

It is instructive to consider the conclusions of the Mb-NO studies in light of previous work on ferrous Mb-CO. The absorption spectrum of the ferrous protein featured three significant peaks in the CO stretching region of the IR [13, 54, 74, 75]. These were assigned to different orientations of the His64 side chain with respect to the ligand, with each orientation contributing a corresponding structural substate with a unique vibrational frequency [76]. Mutation of this distal group subsequently resulted in a different set of potential interactions,

altering the steady-state IR properties of the ligand [77]. 2D-IR studies of Mb-CO showed that a spectral diffusion time constant of 4ps was measured for a CO substate featuring a close interaction with His64 [13]. By contrast, slower spectral diffusion dynamics were observed almost universally in carbonmonoxy Mb mutants with nonpolar distal groups that did not directly contact the ligand [78, 79]. Overall, this suggests that the trends observed using NO to investigate ferric Mb hold across both oxidation states and for both reporter molecules. These combine to indicate that, irrespective of oxidation state, the protein scaffold of Mb appears to give rise to ligand spectral diffusion on a timescale of ~3 ps by virtue of the direct contact. In the event that this contact is not present, the ligand seems subject to slower dynamics, suggesting that the direct contact is the source of the fast motions.

Comparisons of the ferrous and ferric forms suggest that the two show similar structural dynamics, as perhaps would be anticipated if the processes to which the diatomic ligands are sensitive arise from the protein scaffold rather than the Fe centre. This similarity extends to the presence of large pseudostatic contributions to the spectral diffusion parameter as well as to the exponential dynamics. The major difference between the ferric and ferrous forms seems to arise from the presence of multiple conformational states found in ferrous carbonomonoxy proteins versus ferric nitrosylated versions. The reasons for this are not clear and may be based in the relative interaction strengths of the CO and NO with local residues or perhaps to the existence of a larger absorption coefficient for CO making less abundant populations more visible.

Computational methods offer a powerful route to understanding the origins of spectral diffusion processes and further 2D-IR spectroscopy studies were reported on ferrous carbonmonoxy Mb in which comparisons were drawn with molecular dynamics (MD) simulations [13, 80]. These studies addressed ability of MD method to model the separate, but related, issues of spectral diffusion dynamics within a conformational substate [13], as

well as the substate conformational switching time [80]. The latter issue is proposed to be particularly relevant to the functioning of ferrous Mb given the apparent presence of multiple equilibrium geometries [80]. In both studies the T67R/S92D double mutation was used; a species that is of particular interest because the combination promotes peroxidase behaviour in relation to wild-type Mb; potentially allowing insights into the role of dynamics in functionality through comparison of the mutant and wild-type protein. Linear spectra of the double mutant revealed Mb-CO absorption band positions similar to those observed in wild-type Mb [13] though the intensity ratio of the peaks was altered suggesting that mutation influences the preferred histidine conformation.

2D-IR measurements of the spectral diffusion dynamics of these substates showed that similar dynamics were found in both wild-type and mutant Mb forms, suggesting that peroxidase function is not directly ascribable to dramatic changes in the underlying dynamics of the distal cavity. It was reported that these dynamics within the potential well corresponding to a particular conformational substate were well-described by the MD forcefield used [13].

The agreement between MD simulations and timescales measured for conformational switching of the His64 between substates was less accurate however. In the double mutant, such switching was readily observable via exchange 2D-IR spectroscopy experiments [80]. In these experiments, off-diagonal peaks linking the two diagonal peaks corresponding to the individual substates were observed to appear and increase in size as T_w was increased. This was caused by proteins within the ensemble being excited in one conformation and changing to the other during the timescale of the measurement. Thus, the growth of the off-diagonal peaks reflects the switching dynamics between potential surface minima while the change in shape of the 2D-IR peaks reported on the spectral diffusion processes within the individual

wells, as discussed above. For the double mutant, switching timescales were observed in the 100ps range.

Overall, the link between 2D-IR and MD simulations demonstrated that the T67R/S92D mutations give preference to histidine conformations with increased H-bonding contact to the CO. Peroxidase activity was therefore ascribed to alteration of the contact of the distal histidine with the ligand, demonstrating the functional importance of the distal residue and its structural substates to haem function. More recently, new simulation methods have been reported showing impressive agreement with the original experimental results [81]. In particular, this work produced an estimate that the sub-state interconversion time for the wild-type Mb occurs with a characteristic time ~300ps. This timescale for His residue reorientation had been hard to measure experimentally without mutation due to the low intensity of the signals for one of the conformational substates and shows how the link between experiment and theory can contribute to our understanding of these systems [81].

Further insights into the structural origins of the dynamics of Mb were derived from determination of the effects of solvent viscosity and the presence or otherwise of a 'His-tag' [82]. The latter is a biochemical tool consisting of a His_6 unit that is used to assist in protein purification and it was shown that both solvent viscosity and the presence of such a tag can influence the spectral diffusion dynamics observed. Given that the Mb haem pocket is not solvent-accessible and the His_6 tag is located a large distance from the residues which constitute the haem environment, the implication of this work was that the CO probe in the haem pocket is sensitive on some level to non-local dynamics. This conclusion, which was also supported by MD simulations, stands in contrast to earlier assumptions that the haem ligand is a very localised probe [13, 80].

The conclusions reached in studies of Mb appear to hold for other haem proteins. An example is given by neuroglobin (Ngb); an oxygen-transporter within brain tissue. Despite the similar

roles played by Ngb and Mb, their sequence homology is only ~20%, though a high degree of similarity is seen to occur in close proximity to the haem group. 2D-IR was used to investigate ferrous carbonmonoxy Ngb alongside point mutations of Mb that showed similar spectral characteristics [25]. Thus the Mb mutants were used to inform studies of the conformational substates observed for Ngb.

Linear spectra of the L29F form of Mb-CO were seen to bear close resemblance to one structural substate of Ngb-CO; a substate that was assigned to hydrogen bonding interactions between the distal histidine and ligand. Likewise, comparable dynamics were observed between this neuroglobin conformer and L29F over short time-scales. Fast spectral diffusion components on the order of ~2ps were measured in each case, reflecting comparable active site architecture with similar equilibrium fluctuations. Equally, in all cases a substantial static offset suggested similarities in the longer-timescale dynamics [25].

Consistent results were obtained when Ngb structural substates with no histidine contact were studied. In this case the H64V Mb mutant recreated this scenario by substituting the polar histidine for nonpolar valine, eliminating the potential for polar interactions with the ligand. This resulted in the observation of dynamic fluctuations that were substantially slower, with spectral diffusion rates of 11.5ps, relative to the 5.2ps rate in H64V Mb, entirely consistent with the hypothesis that local contacts lead to faster spectral diffusion processes [25].

A final interesting aspect of haemoprotein dynamics, relating to the impact of a disulphide bond upon structural dynamics, was elucidated using Ngb. By comparing the spectral diffusion dynamics of CO bound to Ngb under conditions with and without a key disulphide bond being present, it was shown that disruption of the intramolecular disulphide bond led to increased rates of sampling the different configurations. This implies a significant dynamic impact despite that fact that breaking the bond leads to little impact on the overall structure of the molecule [83].

In addition to determining the origins of fast dynamics, spectral diffusion measurements have provided interesting insights into other aspects of haem protein structure and function. Investigations performed by Khalil *et al*., on the NO transport protein, Nitrophorin4 (NP4), uncovered the potential origin of the protein's pH-triggered, nitric oxide release action [84]. It was observed that, at high and low pHs, NP4 shows two principal conformations. A "closed" form of NP4 is characterised by two hydrogen-bonded loop regions located above the distal haem cavity that gives rise to a tightly-bound, hydrophobic haem pocket. Conversely, the "open" form of NP4 features a solvent accessible, hydrophilic active site, with disordered loop regions that originate from a disrupted hydrogen-bond network. Though each of these conformations is found at either pH, in basic solutions a slightly higher ratio of the open form was observed.

Spectral diffusion in NP4 was reported to originate from a number of sources, over a number of time-scales. It was observed that, at high pH, the slow time component (~100ps) of the open NP4 conformer was significantly faster than that of the low pH protein. With the benefit of MD simulations, this increase in the rate of conformational fluctuation could be attributed to the increased disorder of two loop regions near the distal cavity at high pH. It was proposed that the structural heterogeneity derived from potential loop conformations gave rise to the complex, non-exponential NO release-rate observed in ligand rebinding studies [85]. The association of a reaction mechanism to these regions and their dynamics allows future 2D-IR studies to focus on the significance of particular residues within the loop regions of NP4. With targeted point mutations, combined photolysis, 2D-IR and MD methods have the capacity to identify potential functional differences, determine the impact of mutation on structural dynamics, and isolate the interactions that make specific residues so vital in affecting these behaviours. In addition, it was interesting to note that the work on NP4 also raised the issue of changes in solvation dynamics contributing to the variation in the

spectral diffusion processes [84]. Similar connections had also been drawn in relation to other ferric nitrosylated systems and this opened up another avenue of enquiry in which 2D-IR can be used to ascertain the role of solvent near the haem group of proteins and enzymes [86]. This is explored further in the subsequent section.

Solvent dynamics near haem groups

The 2D-IR studies discussed in the preceding section illustrate the sensitivity of probe ligands to the equilibrium fluctuations of the haem architecture, and how the mode of contact between the distal residue and the ligand is of great significance in determining the dynamics observed. It must also be considered however that the ligand is not restricted to interactions with side-chains in the haem cavity and that the presence of solvent molecules near the haem could perhaps similarly influence ligand dynamics. Such a situation opens up the possibility of 2D-IR spectroscopy making major advances in our understanding of enzymatic processes, as so-called disordered water is hard to detect using crystallography. Nevertheless, the route taken by a substrate between the bulk solvent and the active site could influence the mechanism and efficiency of the enzyme. Furthermore, it is plausible that structural water molecules interacting with haem ligands may play a role in proton transfer or similar steps in an enzymatic process [12]. Examples of such a scenario and the contribution of 2D-IR spectroscopy are discussed below.

The first example considers a 2D-IR investigation of the structural dynamics of the catalase enzyme [86]. This haem-based enzyme is found in almost all aerobically-respiring organisms and is responsible for catalysing the breakdown of hydrogen peroxide into water and molecular oxygen. Interestingly, the structure of the haem pocket of catalase shows some similarity with that of Mb in that it is highly conserved, and features a distal histidine group in close proximity to the haem. As with Mb, the histidine residue of catalase has been strongly linked with protein function. In the case of catalase, the role of this histidine residue

is somewhat different to that of Mb, being thought to facilitate deprotonation of hydrogen peroxide following binding to the ferric Fe centre *en route* to formation of the important oxy-ferryl (Fe(IV)=O) porphyrin radical intermediate called compound I (CpdI) that is central both to the mechanism of catalases and the closely-related peroxidase enzyme family.

In the 2D-IR investigation of catalase (from *Corynebacterium glutamicum*), the use of a NO ligand allowed a study of the enzyme active site in its physiologically-relevant ferric form [86]. As NO is a known inhibitor of catalase *in vivo,* studying this interaction has some biochemical relevance in its own right and it was interesting to note the observation of spectral diffusion dynamics featuring a 3 ps decay timescale (Fig 5a). Comparisons with the results discussed in the last section and the presence of a distal histidine residue in catalase would invite assignment of this timescale to the presence of an interaction of this side chain with the ligand. However, X-ray crystallography studies carried out in parallel with the 2D-IR experiments suggested that a conserved bound water molecule was a more likely source of the interaction (Fig 5b) [86].

This study also observed that catalase showed a remarkably small value of the spectral diffusion parameter at waiting times of 30 ps. This implied that, unlike all haem proteins studied so far, the haem ligand of catalase has the ability to experience the entire continuum of environments available to it in a short time. This was attributed to a dynamically confined active site devoid of, or not influenced by, the slower dynamics found in less constrained systems. It was proposed that the fact that catalase only reacts with hydrogen peroxide molecules and so does not therefore need any flexibility in its pocket architecture to accommodate large substrate molecules could be related to this relative inflexibility. The high efficiency of catalase also points to an active site that is very well suited to its specific purpose. Indeed, the X-ray crystal structure showed that several conserved, bound, water

molecules are located within the catalase distal cavity, linking the local side chains in a well-defined geometry as well as contacting the haem ligand [86].

This conclusion suggested an important role for water in the catalase active site and was not without controversy, not least because the active site of catalase is widely considered to be solvent inaccessible, and water molecules produced in the reaction cycle are thought to be removed from the haem vicinity rapidly during the mechanism [87]. For this reason, the involvement of water in the haem pocket of catalase was explored in more detail via comparative measurements of the ultrafast dynamics observed in water and D_2O solvents [27]. This work exploited the fact that the ν_{NO} mode of the Fe(III)-NO unit occurs at a frequency where it is resolvable in both solvents. Confirmation of the original assignment of a direct interaction between the haem ligand and water molecules was achieved by the detection of solvent isotope dependent vibrational relaxation and spectral diffusion parameters [27]. Further control measurements were reported for Mb where the vibrational relaxation dynamics were shown to be independent of the solvent isotope, as would be anticipated for a direct contact between the haem ligand and the protein backbone rather than water molecules.

In addition to confirming the presence of solvent near the haem centre, these measurements were used to infer something of their dynamics and so to shed light on the notion of solvent accessibility or otherwise of the catalase active site. It was noted that the dynamic timescales observed strongly resembled the water dynamics found in microemulsion micelles, where the encapsulation and small free volume, limited by the molecules making up the structure of the micelle, give rise to restricted motion [88]. These observations correlated well with calculations of free volume within the catalase crystal structure, indicating room for a small pool of disordered water molecules close to the haem, seemingly contradicting the notion of an active site devoid of solvent [27].

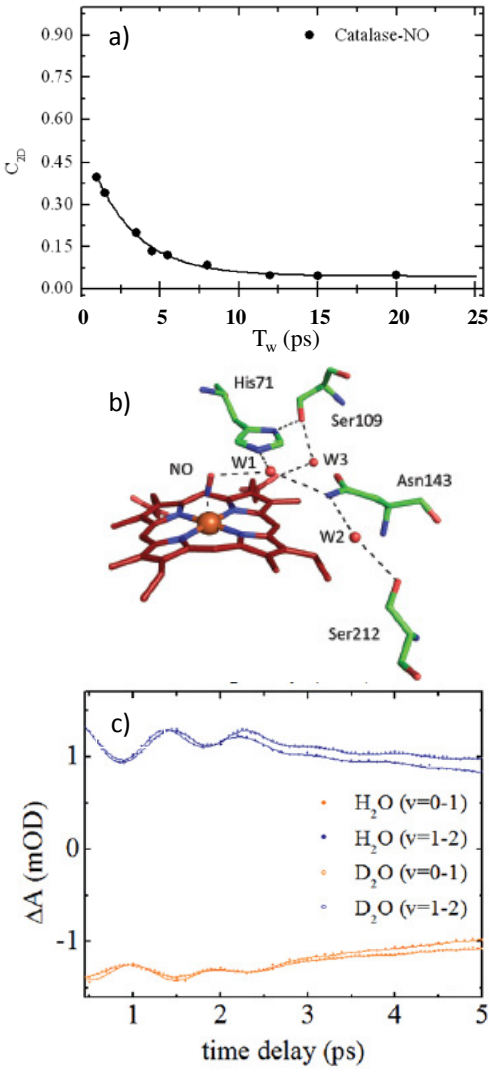

Figure 5: *a) Results of spectral diffusion analysis for catalase-NO from* C. glutamicum*. b) Crystal structure of the same protein (pdb: 4B7F) [86]. c) IR pump-probe spectroscopy results for* C. glutamicum *catalase-NO showing coherent oscillations attributed to water-mediated dynamics [27].*

The mechanistic implications of water interacting with the haem centre in catalase are potentially important and a logical step involved comparing these measurements to similar ones on a peroxidase enzyme [89, 90]. The peroxidases are closely related to the catalases in terms of their mechanisms, both passing initially through CpdI. However, thereafter the mechanisms differ because peroxidases subsequently oxidise two organic substrate molecules and return to the resting Fe(III) state via two single-electron reduction steps while catalase

reacts only with a single H_2O_2 molecule, undergoing a two electron reduction process. Most importantly for this study, the peroxidases are widely accepted to possess solvent accessible active sites [91]. Interestingly, 2D-IR investigations showed that the isotope-sensitive, micelle-like water dynamics observed in catalase remained in peroxidase (from Horseradish), though the spectral diffusion dynamics of the peroxidase did not feature the small pseudostatic contribution of catalase; returning instead to the larger component more typical of other haem proteins [27].

The result of these studies was that, rather than being very different in terms of their water accessibility, these two enzymes were found to be rather similar with both featuring a pool of disordered water molecules near the haem centre, which it was proposed contributes to the proton transfer steps leading to CpdI. Instead, the major difference between the enzyme types was found to lie in the overall flexibility of the peroxidase structure, as determined by the pseudostatic parameter. This suggests a dynamic origin for at least part of the mechanism of these important species and correlates well with the need for peroxidases to accommodate a range of organic substrates in the second step of the reaction; a topic to which we will return in the subsequent section [27].

Finally, it was noted that the pump-probe signal derived from the ν_{NO} mode of nitrosylated catalase featured an oscillatory component with a frequency of ~37 cm^{-1} that is so far unique to catalase, though it may prove simply to be the first such observation of the effect (Fig 5(c)) [27]. Exploiting comparisons with measurements on other haem systems, this was assigned to a coherent phenomenon arising from coupling of the haem ligand stretching vibration to the haem doming vibration mediated by the distal bound water H-bonding structure surrounding the catalase haem [27]. This observation suggests a mechanism whereby the reaction coordinate (stretching of the haem ligand bond) can be directly coupled to protein dynamic motion, reminiscent of the type of ‘promoting vibrations’ that have long been a topic of

debate [92-94]. Given the rapid time-scales of H-bonding fluctuations, 2D-IR is uniquely suited to resolving these dynamics.

Solvent-related dynamics have also been proposed for the previously-discussed NO transport protein, NP4. Although isotope dependent data were not obtained, the dynamics for this system were consistent with a model in which water molecules are trapped in the haem cavity of the closed conformer at high pH. In the low pH regime, where NO release is less favourable, the absence of solvent molecules gives rise to slower structural fluctuations around the NO ligand. In this case, the more rapid structural changes in the haem pocket may result in less successful rebinding dynamics for nitric oxide [84].

The influence of substrate binding on haem pocket dynamics

The remaining factor that may influence the environment or dynamics of a small molecule reporter bound to a haem centre is the presence of a substrate molecule. An example of this referred to above is horseradish peroxidase, which has the capacity to bind and oxidise small aromatic molecules. This is achieved via reduction of the Fe(IV)=O CpdI intermediate that is common to both catalase and peroxidase [12]. A molecular-level explanation has yet to be provided for the reaction mechanism, though it is known that the substrate binds at the edge of the haem pocket via interactions with both the enzyme backbone and the haem propionate group.

A number of studies have utilised benzohydroxamic acid (BHA) as a surrogate substrate within the peroxidase substrate binding site [95-103]. Though it is not the natural substrate of the enzyme, which is not currently known, BHA is one of a range of organic molecules that bind in the substrate position and in fact shows greater affinity for horseradish peroxidase than any other substrate model. Of particular interest is the way in which the protein structure locates the substrate in the correct position to facilitate reaction but also any effect of the

substrate on the local environment. For example, structural investigations of the peroxidase:BHA complex have revealed altered distal cavity interactions near the haem. In particular, crystallography suggests that H-bonding between the distal arginine, R38, and the hydroxamate moiety of BHA stabilises binding of this substrate [100]. More distant effects are also noted through changing of the position of a particular residue side chain that leads to occlusion of the solvent access channel [95].

A direct comparison of the structural dynamics of the peroxidase:BHA complex with those of the substrate-free enzyme was performed for ferric nitrosylated peroxidase [28]. This followed the study that assigned haem pocket dynamics to solvent motion [27] and so compared samples in both aqueous and deuterated water. For nitrosylated peroxidase, the addition of BHA shifted the NO stretching vibrational mode to higher frequency relative to the substrate-free enzyme and the linewidth was significantly reduced. Most strikingly, vibrational lifetime measurements revealed that the isotopic composition of the solvent had no effect on the rate of relaxation. Thus, with substrate bound, the distal cavity appears unable to facilitate solvent-assisted relaxation in the same way as the substrate-free enzyme. In addition to affecting the linewidth of the IR absorption, substrate binding had an impact on the spectral diffusion of the NO ligand; once again, the isotope-dependence of the substrate free enzyme was lost, though the slow dynamics were only weakly affected by complex formation, as indicated by similar pseudostatic values for the aqueous and deuterated systems, respectively.

The conclusion of this study was that, although inhomogeneous broadening of the ν_{NO} mode in substrate-free peroxidase results from direct contact with water molecules, the situation is very different in the substrate complex. Interestingly, the broadening observed in the latter case seems to have its origins in dynamic processes of broadly similar time-scale to the substrate-free enzyme. Presumably, these are defined in both instances by the encompassing

presence of the protein architecture, however it is clear that these dynamics are no longer transmitted to the haem ligand *via* water molecules in the distal cavity in the presence of the substrate, indicating that a chemical, if not dynamic, change to the haem pocket has accompanied binding [28]. Studies of the free volume accessible to unstructured water in substrate-free peroxidase, referred to in the preceding section, suggest that the pool of solvent water within the distal cavity is slightly displaced from the haem group. Thus, it seems that BHA is capable of eliminating isotope effects from the vibrational data by displacing the water molecules that can interact with the haem ligand. This would explain the loss of the isotope dependence and the changes in NO stretching frequency and linewidth upon substrate coordination, which indicate more direct, less inhomogeneous local interactions [103]. Crystal structures of the peroxidase:BHA complex indicate a hydrogen bonding interaction, not only between BHA and the active site ligand, but to the distal residues H42 and R38, entirely consistent with this hypothesis [98, 104, 105].

Prior to the study of the nitrosylated ferric system, 2D-IR studies of ferrous carbonmonoxy horseradish peroxidase also reported changes in the rate of haem ligand structural fluctuations upon BHA binding [106]. It is noteworthy that the absolute spectral diffusion time-scales measured for the ferric peroxidase:BHA complex were in reasonably good agreement with the ferrous system. However, the substrate free enzymes do show some differences. As observed for Mb the infrared absorption spectrum of the ferrous carbonmonoxy enzyme revealed a complex set of haem ligand binding environments, whereas the ferric nitrosylated variant showed only one [28]. In the ferrous form, a significantly higher pK_a is reported for the distal histidine, causing it to display imidazole character at neutral pH. This effect may compete with distal solvent molecules for ligand contact, with different outcomes in each of the haem oxidation states. However the distance to the distal histidine is now thought to be too large for direct contact [107, 108]. Alternatively, the differences may arise from the

greater capacity of nitric oxide to interact with the lone pair of a solvent molecule [61]. While two distinct conformations are detected in substrate-free ferrous peroxidase, BHA binding reduces this to one and so this may well create the same final outcome found in the ferric system [106]. It is clear that, in either case, substrates such as BHA can modify the pre-existing distal H-bond network that connects the active site ligand, distal residues and solvent molecules and that this may be an important part of the mechanistic processes.

The incorporation of substrate binding in haem protein studies can provide further insight to haem protein versatility. 2D-IR measurements have demonstrated how the ultrafast structural fluctuations of cytochrome P450 actively contribute to substrate recognition [23]. Cytochrome P450 binds a plethora of small molecules of varying size and chemical structure [20, 109]. Although the efficiency of hydroxylation varies somewhat between substrates [110-112], only small differences in the X-ray crystal structure of the active site are seen to occur across the various substrate-P450 complexes. Despite this, 2D-IR measurements of substrate-free cytochrome P450 reveal structural heterogeneity with three possible sub-states for the CO mode [23].

The dynamics underpinning the inhomogeneous broadening of these conformational states were found to be slow, resulting in spectral diffusion times from 10-17ps. It was suggested that these could be attributable to either protein side chain motion or perhaps restricted water molecule motion; the interaction of CO with these structural water molecules is thought to vary between substates making the latter a possibility [23]. As was found in ferrous peroxidase, binding of a substrate to cytochrome P450 caused the stabilisation of one particular sub-state. The specific conformation that was stabilised varied for different substrates, though it was notable that the natural substrate, camphor, stabilised a configuration with the lowest frequency CO, while the non-natural substrates each resulted in a single, higher frequency band. Lineshape evolution occurred more rapidly in each of the

complexes upon substrate binding, with an increase in the rate of both the long and short spectral diffusion processes observed. It was further noted that a correlation existed between the rate of long time-scale conformational changes and the substrate dissociation constant, with stronger binding observed in complexes with slower frequency fluctuations. Intuitively, binding of camphor led to the slowest conformation changes, with a time constant of ~370ps. Fluctuations on this time-scale were therefore proposed to relate to binding of the substrate [23]. More rapid conformational sampling was observed for non-natural substrates as result of the lower energy barriers that separate states. In turn, it was proposed that regiospecificity of the cytochrome P450 reaction with camphor could arise from the high energy barriers between states, limiting the transition state pathways [23].

Overall it is clear that, although such studies of enzyme-substrate systems are still in their infancy, trends are beginning to emerge across various sub-families of proteins that suggest that structural dynamics can relate directly to enzymatic mechanistic steps or to the processes such as substrate binding or access that ultimately influence the rate or specificity of the catalyst. Such knowledge would be impossible without the advent of 2D-IR spectral diffusion measurements.

Concluding remarks

In the intervening time since the first measurement of haem protein dynamics, significant progress has been made. It is now possible to use 2D-IR spectroscopy to interrogate protein structural processes ranging from H-bond or disordered water molecule fluctuations to exchange of conformational substates and to discern local from more global dynamic processes. In addition, routes to establishing the role or presence of solvent and the impact of substrate binding have been demonstrated.

In the introduction to the article, it was stated that one of the key aims underlying the application of 2D-IR to heamoprotein systems is to advance our understanding of how dynamics contribute to biologically-relevant processes within the active site. Considering that the experiments reported herein span little more than a decade, much of which has been dedicated to method development, the progress made towards this goal is promising. Having established the origins of the observed dynamics, steps towards correlating them with biological activity have been made with patterns emerging that link faster, 'tighter' dynamics to the binding of 'correct' substrates or which show the existence of dynamical differences between conformations that lead to reaction versus those that do not. Furthermore, observations that could support ideas of promoting vibrations have also arisen from 2D-IR experiments. A key remaining obstacle however is to establish a causative link between the dynamics and motions observed in a 2D-IR experiment on picosecond timescales and the biochemical process itself. It is clear that proteins exhibit dynamics on many timescales and the next goal must be to begin to link these processes together over the full spectrum of time and length scales and then to attribute them directly to biochemical activity. Specific avenues of enquiry can be envisioned such as investigating the manner in which ultrafast fluctuations specifically stimulate secondary structure transformations, or, conversely, the role that fast motions play near transition states that have been achieved following substrate binding or a secondary/tertiary structure change.

To achieve this, it is vital that 2D-IR experiments are applied together with the structural insight that is offered by X-ray crystallography and NMR spectroscopy. In the case of haemoproteins, it is now possible to see how 2D-IR spectroscopy can provide a complementary viewpoint that fits alongside these more established techniques. The ability of 2D-IR to bridge the gap between the timescales accessible by NMR and the limited windows currently open to molecular dynamics simulations will also be an important step in building

up a picture of protein function that covers the full scope of time and distance-dependent effects and their interplay within a single molecule. It is clear from the studies discussed here that progress towards this goal is underway but a still greater library of experiments and systems will be needed. To this end, it is encouraging to see 2D-IR spectrometers moving toward commercial availability and wider accessibility; a step that can only serve to increase the rate of future progress.

Acknowledgements

The authors would like to acknowledge funding from the Engineering and Physical Sciences Research Council, UK, and the Science and Technology Facilities Council to support work featured in this article. The contributions of many co-authors to the research work presented are also gratefully acknowledged.

References

1. M. T. Forrester and M. W. Foster, Free radical biology & medicine, **52**, 1620-1633, (2012)
2. M. Laberge and T. Yonetani, IUBMB Life, **59**, 528-534, (2007)
3. H. Michel, J. Behr, A. Harrenga and A. Kannt, Annual review of biophysics and biomolecular structure, **27**, 329-356, (1998)
4. S. M. Mense and L. Zhang, Cell research, **16**, 681-692, (2006)
5. D. Wendehenne, A. Pugin, D. F. Klessig and J. Durner, Trends in plant science, **6**, 177-183, (2001)
6. K. S. Suslick, *Shape-selective oxidation by metalloporphyrins*, Academic Press, New York, 2000.
7. B. Strandberg, R. E. Dickerson and M. G. Rossmann, Journal of Molecular Biology, **392**, 2-32, (2009)

8. M. S. Smyth and J. H. J. Martin, Molecular Pathology, **53**, 8-14, (2000)

9. C. Göbl and N. Tjandra, Entropy, **14**, 581-598, (2012)

10. L. E. Kay, Journal of magnetic resonance **173**, 193-207, (2005)

11. A. G. I. Palmer, ChemInform, **35**, 3623-3640, (2004)

12. T. L. Poulos, Chemical Reviews, **114**, 3919-3962, (2014)

13. S. Bagchi, B. T. Nebgen, R. F. Loring and M. D. Fayer, Journal of the American Chemical Society, **132**, 18367-18376, (2010)

14. N. Yeung, Y.-W. Lin, Y.-G. Gao, X. Zhao, B. S. Russell, L. Lei, K. D. Miner, H. Robinson and Y. Lu, Nature, **462**, 1079-1082, (2009)

15. J. Y. Zhuang, J. H. Amoroso, R. Kinloch, J. H. Dawson, M. J. Baldwin and B. R. Gibney, Inorganic Chemistry, **43**, 8218-8220, (2004)

16. S. I. Ozaki, M. P. Roach, T. Matsui and Y. Watanabe, Accounts of Chemical Research, **34**, 818-825, (2001)

17. A. M. Lesk and C. Chothia, Journal of molecular biology, **136**, 225-270, (1980)

18. K. G. Welinder, European journal of biochemistry, **151**, 497-504, (1985)

19. M. O. Dayhoff, *Atlas of protein sequence and structure vol 5 supplement 3*, National Biomedical Research, 1979.

20. O. Gotoh, Journal of Biological Chemistry, **267**, 83-90, (1992)

21. W. Doster, D. Beece, S. F. Bowne, E. E. DiIorio, L. Eisenstein, H. Frauenfelder, L. Reinisch, E. Shyamsunder, K. H. Winterhalter and K. T. Yue, Biochemistry, **21**, 4831-4839, (1982)

22. K. A. Henzler-Wildman, M. Lei, V. Thai, S. J. Kerns, M. Karplus and D. Kern, Nature, **450**, 913-916, (2007)

23. M. C. Thielges, J. K. Chung and M. D. Fayer, Journal of the American Chemical Society, **133**, 3995-4004, (2011)

24. J. N. Bandaria, S. Dutta, M. W. Nydegger, W. Rock, A. Kohen and C. M. Cheatum, Proceedings of the National Academy of Sciences of the United States of America, **107**, 17974-17979, (2010)

25. H. Ishikawa, I. J. Finkelstein, S. Kim, K. Kwak, J. K. Chung, K. Wakasugi, A. M. Massari and M. D. Fayer, Proceedings of the National Academy of Sciences of the United States of America, **104**, 16116-16121, (2007)

26. K. Adamczyk, M. Candelaresi, R. Kania, K. Robb, C. Bellota-Antón, G. M. Greetham, M. R. Pollard, M. Towrie, A. W. Parker, P. A. Hoskisson, N. P. Tucker and N. T. Hunt, Physical chemistry chemical physics, **14**, 7411-7419, (2012)

27. K. Adamczyk, N. Simpson, G. M. Greetham, A. Gumiero, M. A. Walsh, M. Towrie, A. W. Parker and N. T. Hunt, Chemical Science, **6**, 505-516, (2015)

28. N. Simpson, K. Adamczyk, G. Hithell, D. J. Shaw, G. M. Greetham, M. Towrie, A. W. Parker and N. T. Hunt, Faraday discussions, (2015)

29. H. Ishikawa, K. Kwak, J. K. Chung, S. Kim and M. D. Fayer, Proceedings of the National Academy of Sciences of the United States of America, **105**, 8619-8624, (2008)

30. S. Bagchi and M. D. Fayer, Biophysical Journal, **98**, 743A-743A, (2010)

31. S. Park and M. D. Fayer, Proceedings of the National Academy of Sciences of the United States of America, **104**, 16731-16738, (2007)

32. D. E. Rosenfeld, K. Kwak, Z. Gengeliczki and M. D. Fayer, Journal of Physical Chemistry B, **114**, 2383-2389, (2010)

33. A. Ghosh, J. Qiu, W. F. DeGrado and R. M. Hochstrasser, Proceedings of the National Academy of Sciences of the United States of America, **108**, 6115-6120, (2011)

34. Y. S. Kim, L. Liu, P. H. Axelsen and R. M. Hochstrasser, Proceedings of the National Academy of Sciences of the United States of America, **106**, 17751-17756, (2009)

35. J. J. Loparo, S. T. Roberts and A. Tokmakoff, Journal of Chemical Physics, **125**, 194522, (2006)

36. J. J. Loparo, S. T. Roberts and A. Tokmakoff, Journal of Chemical Physics, **125**, 194521, (2006)

37. C. J. Fecko, J. J. Loparo, S. T. Roberts and A. Tokmakoff, Journal of Chemical Physics, **122**, 054506, (2005)

38. M. H. Vos, Biochimica et Biophysica Acta - Bioenergetics, **1777**, 15-31, (2008)

39. X. Ye, A. Yu and P. M. Champion, Journal of the American Chemical Society, **128**, 1444-1445, (2006)

40. S. G. Kruglik, B.-K. Yoo, S. Franzen, M. H. Vos, J. L. Martin and M. Negrerie, Proceedings of the National Academy of Sciences of the United States of America, **107**, 13678-13683, (2010)

41. V. Karunakaran, I. Denisov, S. G. Sligar and P. M. Champion, Journal of Physical Chemistry B, **115**, 5665-5677, (2011)

42. V. Karunakaran, A. Benabbas, H. Youn and P. M. Champion, Journal of the American Chemical Society, **133**, 18816-18827, (2011)

43. A. Benabbas, X. Ye, M. Kubo, Z. Y. Zhang, E. M. Maes, W. R. Montfort and P. M. Champion, Journal of the American Chemical Society, **132**, 2811-2820, (2010)

44. M. Kubo, F. Gruia, A. Benabbas, A. Barabanschikov, W. R. Montfort, E. M. Maes and P. M. Champion, Journal of the American Chemical Society, **130**, 9800-9811, (2008)

45. Z. Ganim, H. S. Chung, A. W. Smith, L. P. Deflores, K. C. Jones and A. Tokmakoff, Accounts of Chemical Research, **41**, 432-441, (2008)

46. P. Hamm and M. T. Zanni, *Concepts and Method of 2D Infrared Spectroscopy*, Cambridge University Press, Cambridge, 2011.

47. J. Bredenbeck, J. Helbing, C. Kolano and P. Hamm, Chemphyschem, **8**, 1747-1756, (2007)

48. K. Adamczyk, M. Candelaresi, K. Robb, A. Gumiero, M. A. Walsh, A. W. Parker, P. A. Hoskisson, N. P. Tucker and N. T. Hunt, Measurement Science and Technology, **23**, 062001, (2012)

49. M. D. Fayer, Annual review of physical chemistry, **60**, 21-38, (2009)

50. N. T. Hunt, Chemical Society reviews, **38**, 1837-1848, (2009)

51. Y. S. Kim and R. M. Hochstrasser, Journal of Physical Chemistry B, **113**, 8231-8251, (2009)

52. J. Zheng, K. Kwak and M. D. Fayer, Accounts of Chemical Research, **40**, 75-83, (2007)

53. K. A. Merchant, D. E. Thompson, Q. H. Xu, R. B. Williams, R. F. Loring and M. D. Fayer, Biophysical Journal, **82**, 3277-3288, (2002)

54. K. D. Rector, J. R. Engholm, C. W. Rella, J. R. Hill, D. D. Dlott and M. D. Fayer, Journal of Physical Chemistry A, **103**, 2381-2387, (1999)

55. M. D. Fayer, Annual Review of Physical Chemistry, **52**, 315-356, (2001)

56. R. B. Williams, R. F. Loring and M. D. Fayer, Journal of Physical Chemistry B, **105**, 4068-4071, (2001)

57. Q. H. Xu and M. D. Fayer, Laser Physics, **12**, 1104-1113, (2002)

58. K. A. Merchant, W. G. Noid, D. E. Thompson, R. Akiyama, R. F. Loring and M. D. Fayer, Journal of Physical Chemistry B, **107**, 4-7, (2003)

59. B. L. McClain, I. J. Finkelstein and M. D. Fayer, Journal of the American Chemical Society, **126**, 15702-15710, (2004)

60. C. Rella, K. Rector and A. Kwok, The Journal of Physical Chemistry, **100**, 15620-15629, (1996)

61. T. G. Spiro, A. V. Soldatova and G. Balakrishnan, Coordination chemistry reviews, **257**, 511-527, (2013)

62. E. S. Park, S. S. Andrews, R. B. Hu and S. G. Boxer, Journal of Physical Chemistry B, **103**, 9813, (1999)

63. E. L. Hahn, Physical Review, **80**, 580-594, (1950)

64. M. Maj, Y. Oh, K. Park, J. Lee, K.-W. Kwak and M. Cho, Journal of Chemical Physics, **140**, 235104, (2014)

65. E. S. Park, M. R. Thomas and S. G. Boxer, Journal of the American Chemical Society, **2000**, 12297-12303, (2000)

66. M. Khalil, N. Demirdoven and A. Tokmakoff, Journal of Physical Chemistry A, **107**, 5258-5279, (2003)

67. K. Kwak, S. Park, I. J. Finkelstein and M. D. Fayer, Journal of Chemical Physics, **127**, 124503, (2007)

68. S. T. Roberts, J. J. Loparo and A. Tokmakoff, Journal of Chemical Physics, **125**, 084502, (2006)

69. D. G. Kuroda, J. D. Bauman, J. R. Challa, D. Patel, T. Troxler, K. Das, E. Arnold and R. M. Hochstrasser, Nature Chemistry, **5**, 174-181, (2013)

70. C. Bellota-Antón, J. Munnoch, K. Robb, K. Adamczyk, M. Candelaresi, A. W. Parker, R. Dixon, M. I. Hutchings, N. T. Hunt and N. P. Tucker, Biochemical Society Transactions, **39**, 1293-1298, (2011)

71. U. Flögel, M. W. Merx, a. Godecke, U. K. Decking and J. Schrader, Proceedings of the National Academy of Sciences of the United States of America, **98**, 735-740, (2001)

72. N. T. Hunt, G. M. Greetham, M. Towrie, A. W. Parker and N. P. Tucker, Biochemical Journal, **433**, 459-468, (2011)

73. S. V. Evans and G. D. Brayer, Journal of Molecular Biology, **213**, 885-898, (1990)

74. R. B. Williams, R. F. Loring and M. D. Fayer, The Journal of Physical Chemistry B, **105**, 4068-4071, (2001)

75. M. K. Hong, D. Braunstein, B. R. Cowen, H. Frauenfelder, I. E. Iben, J. R. Mourant, P. Ormos, R. Scholl, A. Schulte and P. J. Steinbach, Biophysical journal, **58**, 429-436, (1990)

76. K. A. Merchant, W. G. Noid, R. Akiyama, I. J. Finkelstein, A. Goun, B. L. Mcclain, R. F. Loring and M. D. Fayer, Journal of the American Chemical Society, **125**, 13804-13818, (2003)

77. R. Maurus, R. Bogumil, N. T. Nguyen, A. G. Mauk and G. Brayer, Biochemical Journal, **74**, 67-74, (1998)

78. K. D. Rector, J. R. Engholm, J. R. Hill, D. J. Myers, H. R., S. G. Boxer, D. D. Dlott and M. D. Fayer, Journal of Physical Chemistry B, **102**, 331-333, (1998)

79. K. D. Rector, C. W. Rella, J. R. Hill, A. S. Kwok, S. G. Sligar, E. Y. T. Chien, D. D. Dlott and M. D. Fayer, Journal of Physical Chemistry B, **101**, 1468-1475, (1997)

80. S. Bagchi, D. G. Thorpe, I. F. Thorpe, G. A. Voth and M. D. Fayer, Journal of Physical Chemistry B, **114**, 17187-17193, (2010)

81. J.-H. Choi, K. Kwak and M. Cho, The Journal of Physical Chemistry. B, **117**, 15462-15478, (2013)

82. M. C. Thielges, J. K. Chung, J. Y. Axup and M. D. Fayer, Biochemistry, **50**, 5799-5805, (2011)

83. H. Ishikawa, S. Kim, K. Kwak, K. Wakasugi and M. D. Fayer, Proceedings of the National Academy of Sciences of the United States of America, **104**, 19309-19314, (2007)

84. M. Cheng, J. F. Brookes, W. R. Montfort and M. Khalil, The Journal of Physical Chemistry. B, **117**, 15804-15811, (2013)

85. M. Kubo, F. Gruia, A. Benabbas, A. Barabanschikov, W. R. Montfort, E. M. Maes and P. M. Champion, Journal of the American Chemical Society, **130**, 9800-9811, (2008)

86. M. Candelaresi, A. Gumiero, K. Adamczyk, K. Robb, C. Bellota-Antón, V. Sangul, J. T. Munnoch, G. M. Greetham, M. Towrie, P. A. Hoskisson, A. W. Parker, N. P. Tucker, M. A. Walsh and N. T. Hunt, Organic and Biomolecular Chemistry, **11**, 7778-7788, (2013)

87. P. Jones, The Journal of biological chemistry, **276**, 1-26, (2001)

88. E. E. Fenn, D. B. Wong, C. H. Giammanco and M. D. Fayer, Journal of Physical Chemistry B, **115**, 11658-11670, (2011)

89. A. M. Azevedo, J. M. S. Cabral and L. P. Fonseca, Biotechnology annual review, **9**, 199-247, (2003)

90. N. C. Veitch, Phytochemistry, **65**, 249-259, (2004)

91. G. I. Berglund, G. H. Carlsson, A. T. Smith, H. Szöke, A. Henriksen and J. Hajdu, Nature, **417**, 463-468, (2002)

92. D. Antoniou and S. D. Schwartz, Journal of Physical Chemistry B, **115**, 15147-15158, (2011)

93. C. R. Pudney, A. Guerriero, N. J. Baxter, L. O. Johannissen, J. P. Waltho, S. Hay and N. S. Scrutton, Journal of the American Chemical Society, **135**, 2512-2517, (2013)

94. L. Masgrau, A. Roujeinikova, L. O. Johannissen, P. Hothi, J. Basran, K. E. Ranaghan, A. J. Mulholland, M. J. Sutcliffe, N. S. Scrutton and D. Leys, Science, **312**, 237-241, (2006)

95. Y.-T. Chang, N. C. Veitch and G. H. Loew, Journal of the American Chemical Society, **7863**, 5168-5178, (1998)

96. H. A. Heering, A. T. Smith and G. Smulevich, Biochemical Journal, **579**, 571-579, (2002)

97. B. D. Howes, H. A. Heering, T. O. Roberts, F. Schneider-Belhadadd, A. T. Smith and G. Smulevich, Biopolymers, **62**, 261-267, (2001)

98. B. D. Howes, J. N. Rodriguez-Lopez, A. T. Smith and G. Smulevich, Biochemistry, **36**, 1532-1543, (1997)

99. M. Laberge, S. Osvath and J. Fidy, Biochemistry, **40**, 9226-9237, (2001)

100. J. S. de Ropp, P. K. Mandal and G. N. La Mar, Biochemistry, **38**, 1077-1086, (1999)

101. J. S. de Ropp, P. K. Mandal, S. L. Brauer and N. La Mar, Journal of the American Chemical Society, **119**, 5400-5407, (1997)

102. G. R. Schonbaum, The Journal of biological chemistry, **248**, 502-511, (1973)

103. B. Zelent, A. D. Kaposi, N. V. Nucci, K. A. Sharp, S. D. Dalosto, W. W. Wright and J. M. Vanderkooi, The Journal of Physical Chemistry B, **108**, 10317-10324, (2004)

104. A. Henriksen, D. J. Schuller, K. Meno, K. G. Welinder, A. T. Smith and M. Gajhede, Biochemistry, **37**, 8054-8060, (1998)

105. A. Gumiero, E. J. Murphy, C. L. Metcalfe, P. C. E. Moody and E. L. Raven, Archives of Biochemistry and Biophysics, **500**, 13-20, (2010)

106. I. J. Finkelstein, H. Ishikawa, S. Kim, A. M. Massari and M. D. Fayer, Proceedings of the National Academy of Sciences of the United States of America, **104**, 2637-2642, (2007)

107. P. Vidossich, G. Florin, M. Alfonso-Prieto, E. Derat, S. Shaik and C. Rovira, Journal of Physical Chemistry B, **114**, 5161-5169, (2010)

108. E. Derat, S. Shaik, C. Rovira, P. Vidossich and M. Alfonso-Prito, Journal of the American Chemical Society, **129**, 6346-6347, (2007)

109. R. Bernhardt, Journal of biotechnology, **124**, 128-145, (2006)

110. R. E. White, M. B. McCarthy, K. D. Egeberg and S. G. Sligar, Archives of Biochemistry and Biophysics, **228**, 492-502, (1984)

111. W. M. Atkins and S. G. Sligar, The Journal of biological chemistry, **263**, 18842-18849, (1988)

112. P. Loida, M. Paulsen, G. Arnold, R. Ornstein and S. G. Sligar, The Journal of biological chemistry, **270**, 5326-5330, (1995)